

BUNDESREPUBLIK DEUTSCHLAND



REC'D 11 JUL 2000

WIPO

PCT

EU

#2

**Prioritätsbescheinigung über die Einreichung
einer Patentanmeldung**

DE 00/1116

Aktenzeichen: 199 31 050.5

Anmeldetag: 6. Juli 1999

Anmelder/Inhaber: Siemens AG, München/DE

Bezeichnung: Verfahren und Vorrichtung zur Sprachverarbeitung

IPC: G 10 L 5/00

Die angehefteten Stücke sind eine richtige und genaue Wiedergabe der ursprünglichen Unterlagen dieser Patentanmeldung.

München, den 27. Juni 2000
Deutsches Patent- und Markenamt
Der Präsident
Im Auftrag

HOIß

**PRIORITY
DOCUMENT**
SUBMITTED OR TRANSMITTED IN
COMPLIANCE WITH RULE 17.1(a) OR (b)

THIS PAGE BLANK (UPTO)

Beschreibung

Verfahren und Vorrichtung zur Sprachverarbeitung

- 5 Die Entwicklung alltagstauglicher Spracherkennungs- und Sprachsteuersysteme stellt seit Jahren eine der Hauptentwicklungslinien der Computertechnik dar. Im Zuge dieser Entwicklung wurden erhebliche Fortschritte erreicht und marktfähige Spracherkennungssysteme etabliert, die sich im praktischen
- 10 Einsatz auch bewähren. Fortgeschrittene Systeme dieser Art sind auch grundsätzlich zur Sprachsteuerung eines Computers bzw. von angeschlossenen Peripheriegeräten geeignet. Einfache Spracherkennungssysteme, die allerdings nur ein relativ geringes Vokabular verarbeiten können, werden auch bereits in
- 15 den Bereichen Konsumelektronik und KFZ-Ausrüstung sowie weiteren Bereichen eingesetzt, in denen eine akustische Steuerung von Geräten aufgrund eines begrenzten Vokabulars möglich und sinnvoll ist.
- 20 Bei Spracherkennungssystemen gibt es in der Regel Werkzeuge, mit denen das vom Spracherkennungssystem zu erkennende Vokabular eingegeben werden kann. In der Regel werden die Wörter oder Äußerungen über eine entsprechende Oberflächen-Software des Computerprogramms in orthographischer Schreibweise einge-
- 5 geben und in die interne Schreibweise des Spracherkennungssystem (meist eine Variante von phonetischer Umschrift (Lautschrift)) automatisch umgesetzt. Bei diesem automatischen oder durch Lexikon-Look-Up unterstützten Konversionsprozeß können Fehler in der phonetischen Umschrift entstehen, die
- 30 aus unzulänglichen Umsetzungsregeln und/oder unvollständigen Lexika herrühren. Da das Spracherkennungssystem seinen Erkennungsprozeß auf der Basis dieser so erzeugten phonetischen Umschrift aufbaut, entstehen durch eine falsche phonetische Umschrift auch Fehler bei der Spracherkennung.

Um eine optimale Performance zu gewährleisten, muß dafür Sorge getragen werden, daß die phonetische Umschrift möglichst lückenlos korrekt ist.

5 Das Problem wurde bisher dadurch gelöst, daß der Benutzer nach Eingabe der orthographischen (korrekten) Schreibweise die vom System erzeugte phonetische Umschrift manuell prüfen konnte. Dies ist für ungeschultes Personal aber in der Regel schwierig. Daher wurden verschiedene Hilfsmittel benutzt, die
10 in der marktgängigen SW angeboten sind:

1. Der Benutzer kann sich für die verschiedenen lautlichen Symbole typische Wörter, in denen solche Symbole enthalten sind, anzeigen lassen und die phonetische Schreibweise manuell korrigieren. Hierbei wird er bei manchen Systemen noch
15 dahingehend unterstützt, daß keine falschen Zeichensequenzen der phonetischen Umschrift verwendet werden können, indem durch die eingesetzte Software nur solche Zeichenketten eingegeben werden können, die eine gültige ASCII-Folge für
20 den verwendeten phonetischen Zeichensatz darstellen.

2. Aus der phonetischen Schreibweise wird mit Hilfe von gängigen Text-to-Speech-Softwaresystemen, also Sprachsyntheseverfahren, die phonetische Umschrift wieder in eine hörbare
25 Sprache umgesetzt. Dies dient dem akustischen Plausibilitätscheck der Phonemkette, die vom System für ein Wort automatisch erzeugt wurde. Dieser Hörtest kann jedoch nur drastische Fehler eliminieren und unterliegt den Unzulänglichkeiten des akustischen Kanals. Des weiteren ist eine Übereinstimmung
30 der verwendeten phonetischen Alphabete der Spracherkennung wie auch der Sprachsynthese zu gewährleisten, was in den wenigsten Fällen gegeben ist.

Der Erfindung liegt daher die Aufgabe zugrunde, ein verbessertes Verfahren bzw. eine Vorrichtung zur Sprachverarbeitung anzugeben, die sich insbesondere durch wesentlich verbesserte Nutzerfreundlichkeit und im Zusammenhang hiermit auch durch erhöhte Genauigkeit und Zuverlässigkeit auszeichnen.

Diese Aufgabe wird hinsichtlich ihres Verfahrensaspektes gelöst durch ein Verfahren mit den Merkmalen des Anspruchs 1 und hinsichtlich ihres Vorrichtungsaspektes durch eine Vorrichtung mit den Merkmalen des Anspruchs 6.

Die Erfindung schließt den wesentlichen Gedanken ein, anstelle der für den sprachwissenschaftlich ungeschulten Benutzer ungewohnten und nur schwer handhabbaren Ausgabe eines in phonetische Umschrift umgesetzten Wortes in dieser phonetischen Umschrift (Lautschrift) eine einfache und zuverlässiger handhabbare Ausgabe vorzusehen. Sie schließt weiter den Gedanken ein, hierzu eine als "pseudo-orthographisch" zu bezeichnende Ausgabeform zu wählen, die vom Benutzer nicht die Kenntnis von Sonderzeichen der phonetischen Umschrift und von deren speziellen Regeln erfordert. Vereinfacht gesagt, erfolgt die Ausgabe der umgesetzten Worte "so, wie man sie spricht".

Diese auch für den Laien leicht verständliche und gut zu handhabende pseudo-orthographische Ausgabe einer in phonetische Umschrift umgesetzten Sprache erfordert einen zusätzlichen Schritt im Sprachverarbeitungsverfahren, nämlich den Schritt der Umsetzung aus der phonetischen Umschrift in diese pseudo-orthographische Darstellung. Dieser zusätzliche Schritt beinhaltet ein Verfahren, bei dem in selbstlernender Weise oder unter Zugriff auf ein vorbestimmtes Regelwerk die phonetischen Einheiten der Worte in einfache graphematische Einheiten von geschriebener Schrift umgesetzt werden. Diese Umsetzung geschieht in einer einfachen und zweckmäßigen Ausführung unter Zugriff auf eine gespeicherte Phonem-Graphem-Zuordnungstabelle, die mindestens mit einem Anfangsbe-

stand an Zuordnungsvorschriften initialisiert und ggfs. im Zuge eines Selbstlernprozesses während der Anwendung des Systems aufgrund zusätzlicher Eingaben durch den Nutzer erweiterbar ist.

5

In einer besonders komfortablen und im Sinne des erwähnten Selbstlernprozesses vorteilhaften Ausführung umfaßt das Verfahren noch einen weiteren Umsetzungsschritt der umgekehrten Umsetzung aus einer (durch den Nutzer bei einer Eingabe zur

10

Korrektur des primären Umsetzungsergebnisses benutzten) pseudo-orthographischen Darstellung in die phonetische Umschrift. Auch bei diesem Schritt kann die erwähnte tabellarische Zuordnung genutzt und ggfs. im Zuge eines Selbstlernprozesses ergänzt und verfeinert werden.

15

Entsprechend den oben angegebenen Verfahrensmerkmalen weist eine Vorrichtung zur Durchführung des vorgeschlagenen Verfahrens neben einer an sich bekannten ersten Umsetzereinheit zur Umsetzung einer orthographischen Eingabe in die phonetische Umschrift eine zweite Umsetzereinheit zur Umsetzung aus der phonetischen Umschrift in die erwähnte pseudo-orthographische Darstellung und eine Ausgabereinheit zur Ausgabe in dieser Darstellungsform auf.

20

25

Für die erwähnte Fortbildung des Verfahrens, die dem Benutzer eine korrigierende Eingabe unter Nutzung der pseudo-orthographischen Darstellung ermöglicht, weist die Vorrichtung eine entsprechende dritte Umsetzereinheit auf.

30

Zur Anwendung der erwähnten Phonem-Graphem-Zuordnungstabelle verfügt die Vorrichtung in einer bevorzugten Ausführungsform über einen entsprechenden Speicher, in dem diese Zuordnungstabelle für die zweite und/oder dritte Umsetzereinheit zugreifbar gehalten wird.

35

Vorteile und Zweckmäßigkeiten der Erfindung ergeben sich im übrigen aus den Unteransprüchen sowie der nachfolgenden Be-

schreibung einer bevorzugten Ausführungsform anhand der Figur.

Die Figur zeigt eine schematische Darstellung einer Sprach-
5 verarbeitungsvorrichtung 1 zur Durchführung des erfindungsge-
mäßigen Verfahrens in einer Ausführungsform in Form eines Funk-
tions-Blockschaltbildes. Die Sprachverarbeitungsvorrichtung 1
umfaßt eine akustische Eingabeeinheit 3, an deren Ausgang ein
vorverarbeiteter Sprachstrom S1 bereitsteht, der einem
10 Eingang einer Spracherkennungseinheit 5 zugeführt wird, die
einen geschriebenen Text S2 ausgibt. Die Spracherkennungsein-
heit 5 umfaßt einen Vokabularspeicher 5a, in dem - in der bei
herkömmlichen Spracherkennungssystemen üblichen phonetischen
Notation - das Vokabular der Spracherkennungseinheit gespei-
15 chert ist.

Der Vokabularspeicher 5a wird laufend durch die Eingabe zu-
sätzlicher Begriffe mittels einer alphanumerischen Eingabe-
einheit 7 aktualisiert, die aus dem orthographischen Eingabe-
20 format in einer ersten Umsetzereinheit 9 in die phonetische
Umschrift (Lautschrift) umgesetzt werden. Ein Lexikonspeicher
11 unterstützt die Umsetzungsprozedur in der ersten Umsetzer-
einheit 9. Zur Prüfung und Korrektur vorgenommener Eingaben
ist eine zweite Umsetzereinheit 13 zur Umsetzung der phoneti-
5 schen Umschrift in eine pseudo-orthographische Darstellung
vorgesehen. Diese wird auf einem Bildschirm 15 für den Benut-
zer angezeigt.

Weiterhin ist eine dritte Umsetzereinheit 17 zur Umsetzung
30 pseudo-orthographischer Eingaben über die alphanumerische
Eingabeeinheit 7 in phonetischer Notation vorgesehen, deren
Ausgang mit dem Vokabularspeicher 5a der Spracherkennungsein-
heit 5 verbunden ist. Der zweiten und dritten Umsetzereinheit
13, 17 ist ein in Form eines Lookup-Table organisierter
35 Zuordnungsspeicher 19 für vorbestimmte Phonem-Graphem-Zuord-
nungen zugeordnet.

Eine durch den Benutzer erfolgte Eingabe eines neuen Begriffes in korrekt orthographischer Notation wird in der ersten Umsetzereinheit 9 in Lautschrift umgesetzt und kann - je nach konkreter organisation des Systems in dieser Form bereits dem Vokabularspeicher 5a zugeführt werden. In jedem Falle wird das in Lautschrift umgesetzte Wort aber der zweiten Umsetzereinheit 13 zugeführt, wo eine weitere Umsetzung in eine pseudo-orthographische Darstellung erfolgt, die auf dem Bildschirm 15 angezeigt wird und den Benutzer dazu veranlaßt, ggfs. über die Eingabeeinheit 7 - nunmehr in der pseudo-orthographischen Darstellung, die auch auf dem Bildschirm erscheint - eine korrigierende Eingabe zu machen oder aber die angezeigte pseudo-orthographische Darstellung zu bestätigen. Die pseudo-orthographische Eingabe wird in der dritten Umsetzereinheit 17 in Lautschrift umgesetzt und nun (erstmalig oder, falls das Wort bereits bei der ersten Eingabe in den Vokabularspeicher 5a übernommen wurden, in einem Korrekturmodus) dem Vokabularspeicher 5a zugeführt. Dessen Inhalt ist damit um eine hinsichtlich der phonetischen Notation geprüfte Vokabel erweitert.

Das oben beschriebene Vorgehen wird nachfolgend an zwei Beispielen erläutert:

1. Beispiel

Über die alphanumerische Eingabeeinheit 7 wird in korrekt orthographischer Schreibweise eingegeben: "Jacques Chirac". In der ersten Umsetzereinheit 9 wird daraus die phonetische Notation gebildet: "sh a xk sh i: rr a xk". Die zweite Umsetzereinheit 13 bildet daraus "sch a k sch i r a k", und in dieser Notation wird der eingegebene Name auf dem Bildschirm 15 angezeigt. Aus dieser Darstellung läßt sich - ohne Kenntnis des in der ersten Umsetzung verwendeten phonetischen Alphabets - erkennen, daß die vom System erzeugte phonetische Notation adäquat ist. Der Nutzer kann das Umsetzungsergebnis

bestätigen, und der neu eingegebene Name gelangt (in phonetischer Notation) in den Vokabularspeicher 5a.

2. Beispiel

5

Über die Eingabeeinheit 7 wird eingegeben "Professional Service". Die erste Umsetzereinheit 9 erzeugt daraus in phonetischer Notation

10 "p r o : f a e s h o n : e : l l s o e r v i : c c : e". Im Ergebnis der weiteren Umsetzung in der zweiten Umsetzereinheit 13 ergibt sich daraus in pseudo-orthographischer Notation: "Profäschonell Sörwieke", und diese Darstellung wird wieder auf dem Bildschirm 15 angezeigt.

15 Der Nutzer erkennt ohne weiteres, daß die vom System erzeugte Lautschrift nicht korrekt sein kann, insofern sie nicht der üblichen Aussprache der eingegebenen Wortkombination entspricht. Der Benutzer wird nun mittels der Eingabeeinheit unter Nutzung der pseudo-orthographischen Notation, die auf
20 dem Schirm dargestellt ist, eine Korrektur vornehmen, und das Korrekturergebnis wird in der dritten Umsetzereinheit 17 aus der pseudo-orthographischen wieder in die phonetische Notation umgesetzt und in dieser in den Vokabularspeicher 5a übernommen. Beim gegebenen Beispiel wird der Benutzer also
5 eingeben "Profäschonnell Sörwis", und entsprechend wird die neue Wortkombination (in phonetischer Notation) im Vokabularspeicher verankert.

Es ist einzusehen, daß das angegebene Verfahren auch mehr-
30 schrittig ausgeführt werden kann, indem nach einer ersten Korrektur durch den Benutzer nochmals eine Umsetzung aus der phonetischen Notation in die pseudo-orthographische und eine weitere Anzeige in dieser Darstellung erfolgt, so daß ggfs. Systemfehler iterativ behoben werden können. Bevorzugt ist
35 hierbei die Anwendung eines - an sich bekannten - selbstlernenden Systems etwa in Gestalt eines neuronalen Netzes, mit dem eine Selbstadaption des Speicherinhaltes des Zuordnungs-

speichers 19 und/oder der Zuordnungsvorschriften des ersten Umsetzungsvorganges (orthographisch - phonetisch) erfolgen kann.

- 5 Die Ausführung der Erfindung ist nicht auf das oben beschriebene Beispiel beschränkt, sondern auch in einer Vielzahl von Abwandlungen möglich, die im Rahmen fachmännischen Handelns liegen.

Patentansprüche

1. Verfahren zur Sprachverarbeitung, bei dem in einem ersten
Umsetzungsschritt eine Umsetzung einer orthographischen
5 Eingabe in eine phonetische Umschrift erfolgt und ein Schritt
der Prüfung und Korrektur des Umsetzungsergebnisses durch den
Nutzer vorgesehen ist,
g e k e n n z e i c h n e t d u r c h
einen zweiten Schritt der Umsetzung aus der phonetischen
10 Umschrift in eine pseudo-orthographische Darstellung und die
Ausgabe in dieser Darstellung.
2. Verfahren nach Anspruch 1,
g e k e n n z e i c h n e t d u r c h
15 einen dritten Schritt der Umsetzung einer in der pseudo-
orthographischen Darstellung erfolgten Eingabe in die phone-
tische Umschrift.
3. Verfahren nach Anspruch 1 oder 2,
20 d a d u r c h g e k e n n z e i c h n e t, daß
der zweite und/oder dritte Umsetzungsschritt eine Umsetzung
phonetischer Worteinheiten in einfache graphematische
Schrift-Einheiten bzw. umgekehrt umfaßt.
4. Verfahren nach Anspruch 3,
5 d a d u r c h g e k e n n z e i c h n e t, daß
der zweite und/oder dritte Umsetzungsschritt unter Zugriff
auf eine gespeicherte Phonem-Graphem-Zuordnungstabelle (19)
ausgeführt wird.
- 30 5. Verfahren nach Anspruch 3 oder 4,
d a d u r c h g e k e n n z e i c h n e t, daß
der zweite und/oder dritte Umsetzungsschritt mittels eines
selbstlernenden Verfahrens, insbesondere unter Nutzung eines
35 neuronalen Netzes zur fortlaufenden Aktualisierung der Pho-
nem-Graphem-Zuordnungstabelle (19), ausgeführt wird.

6. Vorrichtung (1) zur Durchführung des Verfahrens nach einem der vorangehenden Ansprüche, mit einer alphanumerischen Eingabeeinheit (7) und einer eingangsseitig mit dieser verbundenen ersten Umsetzereinheit (9) zur Umsetzung einer orthographischen Eingabe in eine phonetische Umschrift und einer Anzeigeeinheit (15) zur optischen Darstellung eines eingegebenen Wortes,
g e k e n n z e i c h n e t d u r c h
eine zweite Umsetzereinheit (13) zur Umsetzung aus der phonetischen Umschrift in eine pseudo-orthographische Darstellung, die ausgangsseitig mit der Anzeigeeinheit verbunden ist.
7. Vorrichtung nach Anspruch 6,
g e k e n n z e i c h n e t d u r c h
eine dritte Umsetzereinheit (17) zur Umsetzung einer in der pseudo-orthographischen Darstellung erfolgten Eingabe in die phonetische Umschrift.
8. Vorrichtung nach Anspruch 6 oder 7,
d a d u r c h g e k e n n z e i c h n e t, daß
die zweite und/oder dritte Umsetzereinheit (13, 17) mit einem Speicher (19) zur Speicherung einer Phonem-Graphem-Zuordnungstabelle verbunden ist.
9. Vorrichtung nach einem der Ansprüche 6 bis 8,
d a d u r c h g e k e n n z e i c h n e t, daß
die zweite Umsetzereinheit (13) ausgangsseitig mit einem Vokabularspeicher (5a) einer Spracherkennungseinheit (5) verbunden ist.

Zusammenfassung

Verfahren und Vorrichtung zur Sprachverarbeitung

- 5 Verfahren zur Sprachverarbeitung, bei dem in einem ersten Umsetzungsschritt eine Umsetzung einer orthographischen Eingabe in eine phonetische Umschrift erfolgt und ein Schritt der Prüfung und Korrektur des Umsetzungsergebnisses durch den Nutzer vorgesehen ist, mit einem zweiten Schritt der Umsetzung aus der phonetischen Umschrift in eine pseudo-orthographische Darstellung und die Ausgabe in dieser Darstellung.
- 10

